

**WestminsterResearch**

<http://www.westminster.ac.uk/westminsterresearch>

**Interpretability and Complexity of Design in the Creation of Fuzzy Logic Systems - A User Study**

**Razak, T.R., Garibaldi, J.M., Wagner, C., Pourabdollah, A. and Soria, D.**

This is a copy of the author's accepted version of a paper subsequently published in the proceedings of the 2018 IEEE Symposium Series on Computational Intelligence. Bengaluru, India, 18-21 Nov 2018.

The final published version will be available online at:

<https://dx.doi.org/10.1109/SSCI.2018.8628924>

© 2018 IEEE . Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

---

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

---

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch: (<http://westminsterresearch.wmin.ac.uk/>).

In case of abuse or copyright appearing without permission e-mail [repository@westminster.ac.uk](mailto:repository@westminster.ac.uk)

# Interpretability and Complexity of Design in the Creation of Fuzzy Logic Systems — A User Study

Tajul Rosli Razak<sup>1,4</sup>, Jonathan M. Garibaldi<sup>1</sup>, Christian Wagner<sup>1</sup>, Amir Pourabdollah<sup>2</sup> and Daniele Soria<sup>3</sup>

<sup>1</sup>Intelligent Modelling & Analysis (IMA) and Laboratory for Uncertainty in Data and Decision Making (LUCID) Groups, School of Computer Science, University of Nottingham, United Kingdom.

<sup>2</sup>School of Science and Technology, Nottingham Trent University, United Kingdom.

<sup>3</sup>Department of Computer Science, University of Westminster, United Kingdom.

<sup>4</sup>Faculty of Computer and Mathematical Science, Universiti Teknologi MARA, Perlis, Malaysia.

Email: <sup>1</sup>{tajul.razak,jon.garibaldi,christian.wagner}@nottingham.ac.uk, <sup>2</sup>amir.pourabdollah@ntu.ac.uk,

<sup>3</sup>d.soria@westminster.ac.uk, <sup>4</sup>tajulrosli@perlis.uitm.edu.my

**Abstract**—In recent years, researchers have become increasingly more interested in designing an interpretable Fuzzy Logic System (FLS). Many studies have claimed that reducing the complexity of FLSs can lead to improved model interpretability. That is, reducing the number of rules tends to reduce the complexity of FLSs, thus improving their interpretability. However, none of these studies have considered interpretability and complexity from human perspectives. Since interpretability is of a subjective nature, it is essential to see how people perceive interpretability and complexity particularly in relation to creating FLSs. Therefore, in this paper we have investigated this issue using an initial user study. This is the first time that a user study has been used to assess the interpretability and complexity of designs in relation to creating FLSs. The user study involved a range of expert practitioners in FLSs and received a diverse set of answers. We are interested to see whether, from the perspectives of people, FLSs are necessarily *more interpretable* when they are *less complex* in terms of their design. Although the initial user study is based on small samples (i.e., 25 participants), nevertheless this research provides initial insight into this issue that motivates our future research.

**Index Terms**—Fuzzy Logic Systems, Interpretability, Complexity of design, User study

## I. INTRODUCTION

Interpretability is acknowledged as one of the most appreciated advantages of Fuzzy Logic Systems (FLSs) [1], particularly in relation to applications such as knowledge extraction and decision support [2], [3]. Specifically, one of the most important motivations for using FLSs for system modelling is that a fuzzy system uses linguistic variables and rules [4] that are easy to understand. However, this rule based structure also poses significant challenges, including the *curse of dimensionality*, where the number of required rules commonly and model complexity increase exponentially with the number of input variables [5], [6], thus potentially reducing the transparency and interpretability of FLSs. From these challenges, we can summarise the relationship between the number of rules, complexity and interpretability in FLSs as follows:

↑ *Number of rules* (FLSs), ↑ *Complexity* (FLSs) (1)

↑ *Complexity* (FLSs), ↓ *Interpretability* (FLSs) (2)

where ↑ denotes increasing and ↓ indicates decreasing.

In exploring this problem, several methods have been proposed for optimising the size of the rule-base in FLSs, such as rule selection [7], feature selection [8], rule interpolation [9], singular-value decomposition-QR [10], evolutionary algorithm [11], fuzzy similarity measure [6], rule learning [12] and hierarchical fuzzy systems (HFSs) [13]. Many of these approaches are designed to overcome the *curse of dimensionality* – thus reducing the model complexity of FLSs [14]. Moreover, by reducing model complexity, we may also find a way to improve model interpretability, i.e., a system with a smaller number of rules requires a smaller effort to be interpreted [15]. However, the following questions arise: (i) “Is it always true that by reducing the complexity of FLSs, we can improve the interpretability of FLSs?”; and (ii) “Does an FLS with fewer rules always have better interpretability compared to an FLS with more rules?”.

In [16] Ishibuchi *et al.* demonstrated, using a simple test problems, that the minimisation of complexity does not always lead to maximisation of interpretability, particularly in the design of the fuzzy rule-based system. However, the problem is that interpretability is a very difficult concept, because of its subjective nature. Indeed, evaluating how different people perceive interpretability is not a trivial matter. Therefore, we believe that it is important to analyse how people perceive interpretability and complexity of design when it comes to the creation of FLSs. In other words, from a human perspectives, “does the minimisation of FLS complexity of design lead to the maximisation of FLS interpretability?”. To the authors knowledge, this intriguing question has never been addressed through the use of a user study.

This paper introduces an initial user study to assess both interpretability and complexity of design in relation to creating FLSs. While the number of participants in the study is still limited, the paper puts forward a valuable first step both in terms of using user studies in respect to exploring the properties of FLSs, and in addressing the key question of complexity vs interpretability.

The rest of this paper is organised as follows. Section II discusses the background to interpretability, complexity, HFSs

and user studies. This is followed by Section III that introduces a user study based approach to assess the interpretability and complexity of the design of FLSs and also HFSs. In Section IV, the process to investigate the relationship between interpretability and complexity of design is carefully explained. Finally, Section V presents the conclusions and future works.

## II. BACKGROUND

In this section, we briefly provide background in respect to interpretability, complexity, HFSs and user studies.

### A. Interpretability

Interpretability refers to the capability of FLSs to express the behavior of the system in an understandable way [17]. In FLSs, fuzzy sets are used to describe the domains of values of certain variables. As in human thinking, linguistic terms can be used for this purpose. This property makes fuzzy systems rather unique among other modelling systems because while maintaining some intuitive conditions about the collection of fuzzy sets, they possess the ability to be easily interpreted, even for layman users [18]. Nevertheless, while fuzzy sets can generally be used to model perceptions, this does not always imply their straightforward interpretation in natural language. This is because interpretability is a subjective property that depends on several factors, mainly the model structure, the number of input variables, the number of fuzzy rules, the number of linguistic terms, and the shape of the fuzzy sets [17] - as well as of course the background and context of the individual.

There is substantial research on interpretability measurement that proposes interpretability indices for FLSs [19]–[24]. However, the choice of an appropriate interpretability measurement is still an open discussion.

### B. Complexity

In FLSs, complexity is expressed by the number of rules, variables, and fuzzy terms. For instance, an FLS with fewer rules is easier to configure and requires less memory and execution time than an FLS with more rules. Thus, in the literature, most studies have used indices, i.e., the number of rules, to measure the complexity of FLSs [6], [25]–[27].

The complexity in FLSs grows exponentially with their input variables [6], which is also known as *complexity explosion*. Model complexity can have a significant impact on the accuracy, efficiency and interpretability of FLSs [28]. For example, as shown in (2), increasing the complexity in FLSs may reduce their interpretability.

While complexity is commonly measured by the number of rules in FLSs, we believe that the topologies of the FLSs, i.e. flat, serial and parallel, also have an impact on their complexity.

### C. Interpretability-Complexity Tradeoff

As discussed earlier, the interpretability of fuzzy systems is still being debated among researchers due to its subjective

nature. Thus far, complexity is often used as an indirect measurement of interpretability in FLSs. Furthermore, several researchers claim that the reduction of complexity in a system can lead to better interpretability of the fuzzy system [6], [29], [30].

However, Ishibuchi *et al.* [16] illustrate that complexity minimisation does not always lead to interpretability maximisation using a test problem, particularly in pattern classification problems. Going beyond this, since interpretability is subjective in nature, we believe that this interpretability-complexity tradeoff should also be considered from the human perspective, i.e., “how do people actually perceive interpretability and complexity of design in relation to creating FLSs?”. To the authors’ knowledge, no one has conducted a study with people to explore the interpretability and complexity of design when creating FLSs.

### D. Hierarchical Fuzzy Systems

HFSs were introduced by Raju *et al.* [13] as an approach to overcome the *curse of dimensionality* arising in conventional FLSs. In HFSs, the original FLSs (e.g., as in Fig. 2) are decomposed into a series of low-dimensional FLSs—fuzzy logic subsystems (e.g., as shown in Figs. 3 and 4). Moreover, the rules in HFSs commonly have antecedents with fewer variables than the rules in FLSs with equivalent function, since the number of input variables of each subsystem is lower [31]. Thus, HFSs tend to reduce rule explosion, thus minimising complexity, and improving model interpretability.

However, it is not clear how users may perceive the interpretability and complexity of design particularly in HFSs that have multiple subsystems, layers and varied topologies.

### E. User Studies

A user study is a method that can be used to conduct design research. Examples of user studies in the context of FLSs include that of Balazs and Koczy [18], who conducted interviews to ask users to define fuzzy sets. In their study, they investigate fuzzy sets by analysing how users interpret the meaning of ‘hot’. Based on user-defined linguistic terms, fuzzy rules and rule bases can be constructed easily, and the paper puts forward that this can lead to a complexity reduction.

Mencar and Fanelli [32] conducted a survey that attempted to provide a complete presentation of interpretability constraints in the literature. The survey aims (i) to give a homogeneous description of all interpretability constraints; (ii) to provide a critical review of such constraints; and (iii) to identify potentially different meanings of interpretability. Alonso *et al.* in [33] evaluate the most used interpretability indices with a user study (in the form of a web poll) to get useful information regarding interpretability assessment. The results show that a fuzzy index is easily adapted to the context of each problem as well as the users’ quality criteria.

In this paper, a user study is conducted to assess how users perceive the interpretability and complexity of FLS designs in relation to creating FLSs.

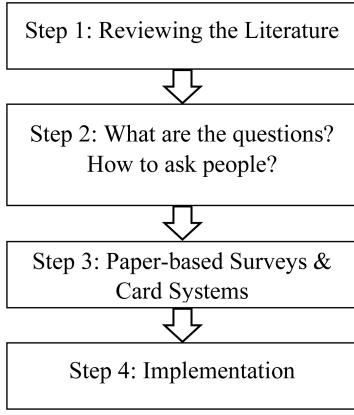


Fig. 1. Four steps towards developing a user study

TABLE I  
IRIS SYSTEMS

Sets	Systems	Number of inputs	Number of MFs	Number of rules
Set MF-2	FLS	4	2	16
	HFS-P	4	2	12
	HFS-S	4	2	12
Set MF-3	FLS	4	3	81
	HFS-P	4	3	27
	HFS-S	4	3	27

### III. THE USER STUDY APPROACH

Setting it apart from the existing research on interpretability, we look at the interpretability and complexity of design in relation to creating FLSs using a user study.

Specifically, we propose an initial user study to assess the interpretability and complexity of the designs in relation to creating FLSs and HFSs. Fig. 1 shows the steps involved in developing this user study:

1) *Step 1 - Reviewing the Literature*: We begin by carefully reviewing the literature that focuses on the problems and issues in relation to interpretability and complexity.

2) *Step 2 - What are the questions? How to ask people?*: We have decided to use a simple Iris flower classification as an example in our questionnaire. The Iris flower consists of four attributes as input features: *sepal length*, *sepal width*, *petal length* and *petal width*. Three classes of Iris flowers are output features: *Setosa*, *Versicolor* and *Virginica*.

In this paper, three Iris systems, are introduced (flat FLS (FLS), Parallel HFS (HFS-P) and Serial HFS (HFS-S)) as can be seen in topologies Fig. 2, Fig. 3 and Fig. 4, respectively. They are classified into two configurations: three systems with two membership functions (MFs), referred to as Set MF-2, and three systems with three MFs, referred to as Set MF-3. The description of these Iris systems can be seen in Table I.

3) *Step 3 - Paper-based Surveys and Card Systems*: As it is challenging to present Set MF-2 and Set MF-3 in a paper-based survey, we decided that each of the Iris systems in Set MF-2 and Set MF-3 should be printed on an A4 card – called a *card system*. The topology, membership functions and rule

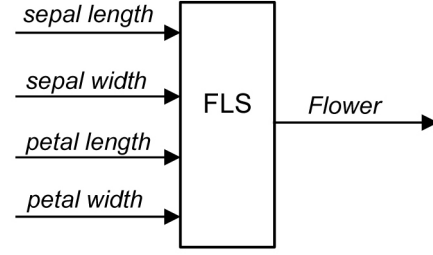


Fig. 2. Iris topology: FLS

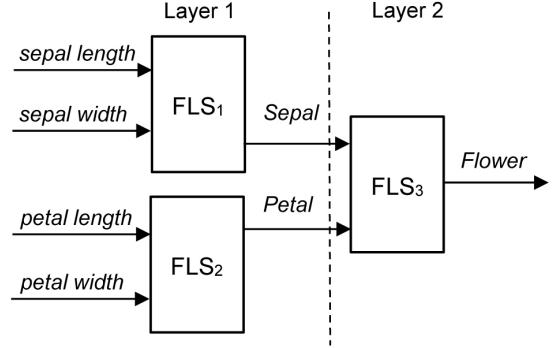


Fig. 3. Iris topology: Parallel HFS

set of each system in Set MF-2 and Set MF-3 were printed on the *card system* (as can be seen in Fig. 5).

In their study, Alonso *et al.* [33] found that, when it comes to interpretability, participants will feel much more confident in giving rankings rather than giving numerical values. By following their method, in the design of questions we decided to ask the participants to provide a rank for the interpretability and complexity of the design of each system in Set MF-2 and Set MF-3.

4) *Step 4 - Implementation*: We carried out this paper-based survey at the IEEE International Conference on Fuzzy Systems 2017, Naples, Italy. For example, as shown in Fig. 6, the participants were asked to rank the three Iris systems on the *card systems* (namely FLS, HFS-P and HFS-S cards) based on

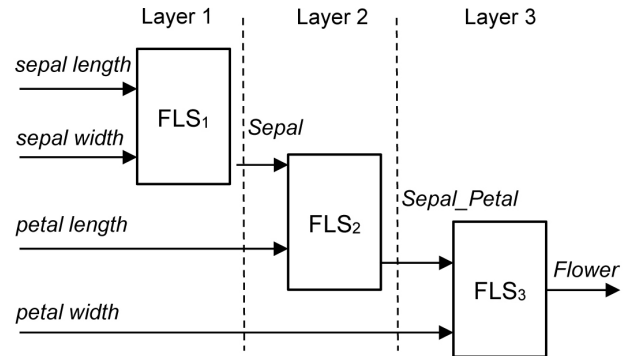


Fig. 4. Iris topology: Serial HFS

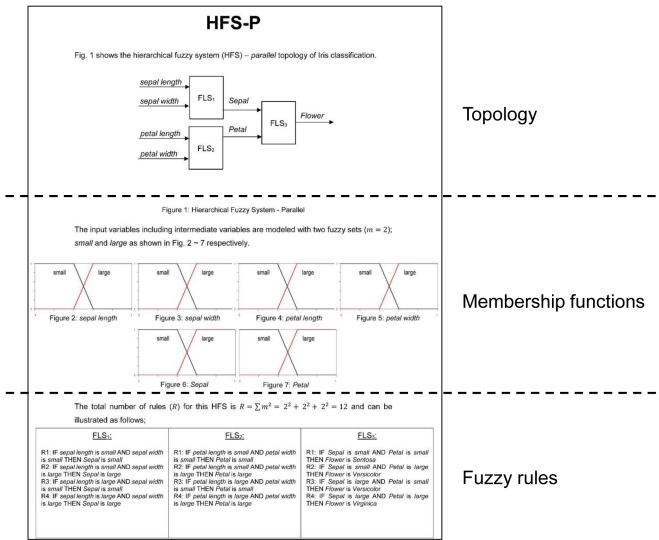


Fig. 5. Card system: HFS-P

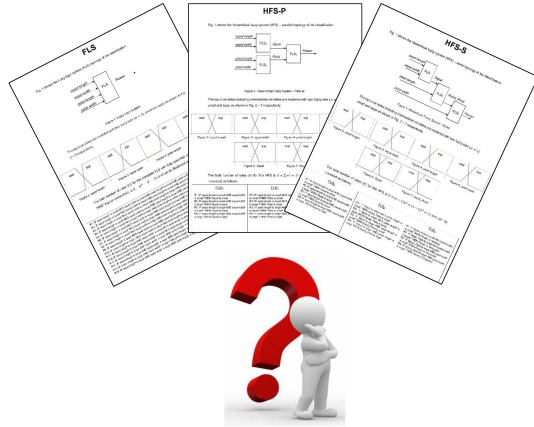


Fig. 6. Participants rank FLS, HFS-P and HFS-S based on the interpretability and complexity of design in relation to creating the Iris systems

their view of the interpretability and complexity of the designs. In terms of the ranking of interpretability, the participants *may* provide rank 1 to indicate the most interpretable of the Iris systems, and rank 3 to indicate the least interpretable of the Iris systems. In contrast, in terms of ranking complexity of design, the participants *may* provide rank 1 to indicate the least complex to design, and rank 3 for the most complex.

#### IV. EXPERIMENTS AND RESULTS

The aim of this study is to investigate the relationship between interpretability and complexity of different FLS designs when creating various Iris classification systems, as discussed in Section III. The analysis process consists of three main steps: (i) descriptive analysis of the survey data; (ii) analysis of individual differences in perceived interpretability and complexity of each of the FLSs – MF-2 and MF-3; and

TABLE II  
FREQUENCY OF THE INTERPRETABILITY RANKINGS IN SET MF-2 AND SET MF-3

Iris systems	Rank						Total	
	1		2		3		Count	(%)
	Count	(%)	Count	(%)	Count	(%)		
Set MF-2:								
FLS	5	20	8	32	12	48	25	100
HFS-P	19	76	4	16	2	8	25	100
HFS-S	2	8	9	36	14	56	25	100
Set MF-3:								
FLS	5	20	5	20	15	60	25	100
HFS-P	18	72	4	16	3	12	25	100
HFS-S	4	16	11	44	10	40	25	100

(iii) statistical analysis of the interpretability and complexity of design data.

#### A. Descriptive Analysis of the Survey Data

We received responses from 25 participants at Fuzz-IEEE 2017 who answered this survey. From this data, we first summarize and describe the survey data in a meaningful way using basic frequency tables.

Table II shows the frequency of the interpretability rankings given by the participants for the Iris systems in Set MF-2 and Set MF-3. For the case of Set MF-2, most of the participants found the HFS-P to be *more interpretable* than the FLS and HFS-S (i.e., rank 1 indicates the most interpretable of the Iris systems). The results show that 76% of the participants chose HFS-P, 20% chose FLS, and 8% chose HFS-S as *more interpretable*. Further, for the case of Set MF-3, it can be seen that most of the participants found that the HFS-P is *more interpretable* than the FLS and HFS-S. The results show that 72% of participants chose HFS-P, 20% chose FLS, and 16% chose HFS-S as *more interpretable*.

On the other hand, Table III shows the frequency of the complexity of design rankings given by the participants in Set MF-2 and MF-3. Here, for the case of Set MF-2, most of the participants chose FLS as *less complex* to design than HFS-P and HFS-S (i.e., rank 1 indicates the least complex to design). The results show that 48% of the participants chose FLS, 40% for chose HFS-P, and 24% for chose HFS-S as *less complex* in terms of the design of the FLS based Iris classification system. Meanwhile, for the case of Set MF-3, the participants chose FLS, HFS-P and HFS-S as *equally complex*. The results show that 40% of participants chose FLS, 40% chose HFS-P, and 40% chose HFS-S as *equally complex* in terms of the design.

#### B. Analysis of Individual Differences in Interpretability and Complexity of Design for Each System

In this Section, we focus on the following questions, “Are the FLSs that participants ranked as *more interpretable*, those that they ranked as the *less complex* to design?” or indeed “Are the FLSs that participants ranked as *more interpretable*, those that they ranked as the *more complex* to design?”. To explore this, Figs. 7, 8, 9, 10, 11 and 12 present the distribution of

TABLE III  
FREQUENCY OF THE COMPLEXITY RANKINGS IN SET MF-2 AND SET MF-3

Iris systems	Rank						Total	
	1		2		3			
	Count	(%)	Count	(%)	Count	(%)	Count	(%)
Set MF-2:								
FLS	12	48	1	4	12	48	25	100
HFS-P	10	40	10	40	5	20	25	100
HFS-S	6	24	7	28	12	48	25	100
Set MF-3:								
FLS	10	40	2	8	13	52	25	100
HFS-P	10	40	10	40	5	20	25	100
HFS-S	10	40	6	24	9	36	25	100

		Complexity of design			
		Less $\longleftrightarrow$ More			Rank
		1	2	3	
Interpretability	More $\updownarrow$ Less	1	3	0	2
		2	7	0	1
		3	2	1	9

Fig. 7. Interpretability vs complexity for FLS in Set MF-2

the participants answers in relation to the interpretability and complexity of the design of each Iris system.

Figs. 7, 8 and 9 present the distribution of participants answers in relation to the interpretability and complexity of the designs for creating FLS, HFS-P and HFS-S respectively, in Set MF-2. It is apparent from these figures that there is substantial diversity in terms of the answers in respect to the interpretability and complexity of the design/creation of FLS, HFS-P and HFS-S. Overall, in the case of FLS, the most salient relationship from Fig. 7 shows that nine participants (36%) ranked FLS as *less interpretable* and *more complex* to design. Similar to FLS, the most salient relationship from Fig. 9 shows that nine participants (36%) ranked HFS-S as *less interpretable* and *more complex* to design. However, for the case of HFS-P, the most salient relationship from Fig. 8 shows that eight participants (32%) ranked HFS-P as *more interpretable* and *less complex* to design.

		Complexity of design			
		Less $\longleftrightarrow$ More			Rank
		1	2	3	
Interpretability	More $\updownarrow$ Less	1	8	7	4
		2	1	3	0
		3	1	0	1

Fig. 8. Interpretability vs complexity for HFS-P in Set MF-2

		Complexity of design			
		Less $\longleftrightarrow$ More			Rank
		1	2	3	
Interpretability	More $\updownarrow$ Less	1	1	0	1
		2	3	4	2
		3	2	3	9

Fig. 9. Interpretability vs complexity for HFS-S in Set MF-2

		Complexity of design			
		Less $\longleftrightarrow$ More			Rank
		1	2	3	
Interpretability	More $\updownarrow$ Less	1	3	1	1
		2	3	1	1
		3	4	0	11

Fig. 10. Interpretability vs complexity for FLS in Set MF-3

Meanwhile, for the case of Set MF-3, Figs. 10, 11 and 12 present a distribution of participant answers ranking the interpretability and complexity of the designs for creating FLS, HFS-P and HFS-S respectively. As shown in these figures, there is also substantial difference in the answers in relation to the interpretability and complexity of design for creating FLS, HFS-P and HFS-S. Likewise, the relationship in Set MF-3 show the same pattern as in Set MF-2. The pattern indicates that the most meaningful relationship from Fig. 10 identified that eleven participants (44%) ranked FLS as *less interpretable* and *more complex* to design. Similar to FLS, the most important relationship from Fig.12 shows that six participants (24%) ranked HFS-S as *less interpretable* and *more complex* to design. However, in the case of HFS-P, the most important relationship from Fig. 11 shows that eight participants (32%) ranked HFS-P as *more interpretable* and *less complex* to design.

Although only the most important relationships (based on the highest participant answers) were discussed for each system, there were also other relationships between interpretability

		Complexity of design			
		Less $\longleftrightarrow$ More			Rank
		1	2	3	
Interpretability	More $\updownarrow$ Less	1	8	7	3
		2	1	3	0
		3	1	0	2

Fig. 11. Interpretability vs complexity for HFS-P in Set MF-3

		Complexity of design			
		Less ← → More			
		Rank	1	2	3
Interpretability	More ↑ ↓ Less	1	2	1	1
		2	5	4	2
		3	3	1	6

Fig. 12. Interpretability vs complexity for HFS-S in Set MF-3

ity and complexity of design in relation to the creation of FLS, HFS-P and HFS-S for both Set MF-2 and Set MF-3. For instance, some participants ranked a *less interpretable* system as *less complex* to design. Also, some participants ranked a *less interpretable* system as *more complex* to design as can be seen in Figs. 7, 8, 9, 10, 11 and 12. Thus, while weaker, these relationships also should be taken into account in order to investigate the significance of the relationship between interpretability and complexity of designs when creating FLSs.

### C. Statistical Analysis

We have investigated the significance of the relationships between interpretability and complexity of design for creating Iris systems that have been captured in Section IV-B, using Pearson's correlation coefficient ( $r$ ). The latter statistically establishes how influential a relationship is between two variables where  $+1$  indicates a strong positive relationship and  $-1$  indicates a strong negative relationship [34].

Table IV provides the summary of the statistics for interpretability and complexity of design for Iris systems in Set MF-2 and Set MF-3 that were computed using Pearson's correlation coefficients. First, results indicate that a positive correlation was found between interpretability and complexity of design in relation to creating FLS, HFS-P and HFS-S for both Set MF-2 and Set MF-3. However, as shown in Table IV, although FLS, HFS-P and HFS-S in Set MF-2 and Set MF-3 gave positive correlations ( $r$ ) between interpretability and complexity of design, only FLS shows statistical significance of this positive correlation ( $r$ ). The computed  $p$ -value for FLS is *less than* 0.05 in both Set MF-2 and Set MF-3. The positive correlations ( $r$ ) in HFS-P ( $r = 0.052$ ,  $r = 0.231$ ) and HFS-S ( $r = 0.316$ ,  $r = 0.275$ ) in Set MF-2 and Set MF-3 respectively, are both not statistically significant.

TABLE IV  
PEARSON CORRELATION COEFFICIENT ( $r$ ) BETWEEN INTERPRETABILITY AND COMPLEXITY OF DESIGN FOR IRIS SYSTEMS IN SET MF-2 AND SET MF-3

Iris systems	Set MF-2		Set MF-3	
	$r$	$p$ -value	$r$	$p$ -value
FLS	0.421	0.036	0.410	0.042
HFS-P	0.052	0.804	0.231	0.266
HFS-S	0.316	0.123	0.275	0.184

### D. Discussion

The study in this paper was conducted as an initial exploration of the relationship between FLS interpretability and complexity of design. It was focused specifically on FLSs designed to implement the Iris classification system, and consisted of three main steps.

For the first step, we presented the survey data, which consists of the interpretability and complexity of designs ranked by participants, as shown in Tables II and III. We found that the majority of the participants chose HFS-P as *more interpretable* than FLS and HFS-S for both Set MF-2 and Set MF-3. In the case of complexity of design, we also found that HFS-P was chosen by most of the participants as *less complex* to design when compared to FLS and HFS-S in Set MF-2. Surprisingly, in Set MF-3, participants ranked FLS, HFS-P and HFS-S as *equally complex* to design. It seems possible that these results are due to different topologies, the number of membership functions, and also the number of rules in each system for both Set MF-2 and Set MF-3, which influence the participants in their ranking of the interpretability and complexity of the designs.

For the second step, an analysis of the individual differences in the interpretability and design complexity rankings for each system was performed as captured in Figs. 7, 8, 9, 10, 11 and 12. We found considerable diversity of relationships between interpretability and complexity of design as given by the participants. For example, some participants stated that the *less interpretable* a system was, the *more complex* it was to design. However, some participants also stated that the *less interpretable* a system was, the *less complex* it was to design. Some possible explanations of this finding include: (i) participants ranked the system as *more interpretable* because they perceived it as being more intuitive and easy to understand for both Set MF-2 and Set MF-3; (ii) the difference in the number of rules between FLS and HFS in Set MF-3 is larger than Set MF-2 – therefore, participants feel that it will be challenging to design an FLS model; and (iii) some participants find it challenging to design HFS-P and HFS-S as they are composed of multiple subsystems and layers.

In order to examine the significance of the relationship between interpretability and complexity of design, the Pearson correlation was established for the third step. A key outcome from the evaluation is that a positive correlation ( $r$ ) between interpretability and complexity of design was found for FLS, HFS-P and HFS-S in both Set MF-2 and Set MF-3, as shown in Table IV. That is, it indicates that the *more interpretable* a system is, the *less complex* it is to design the system, and, also, the *less interpretable* a system is, the *more complex* it is to design the system – at least in the context of the Iris systems considered. However, this positive relationship is statistically significant only for FLS and not for HFS-P and HFS-S for both Set MF-2, and Set MF-3. It is hard to interpret the cases for HFS-P and HFS-S because the statistical power is low due to the small sample size. Nevertheless, it suggests that the effect is weak for the positive correlation for HFS-P and HFS-S.

## V. CONCLUSIONS

In this paper, an initial user study has been used to explore the relationship between interpretability and complexity of design when creating FLSs. This has never been addressed by researchers in this field. Although the current study is based on a small sample of participants and focuses on the particular case of FLSs designed for Iris classification, the findings suggest that, in creating fuzzy systems, it does not always hold true that the *less complex* the design of a system is, the *more interpretable* the system is. This seems especially true when a FLS is hierarchical. We also observed that the topologies of the FLSs (e.g. ‘flat’, ‘parallel’ and ‘serial’) have a significant influence on the interpretability and complexity of the design of FLSs.

While this initial study provides intriguing insight on the relationship between interpretability and complexity as perceived by people (users, participants, etc.), more research on this topic needs to be undertaken, with a larger sample size, in order to develop a fine-grained understanding of the relationship between interpretability and complexity of design when creating FLSs.

## ACKNOWLEDGMENT

We would like to thank all participants for giving up their time for this study. Their answers and detailed comments were invaluable for the completion of this work.

We gratefully acknowledge the support of the Universiti Teknologi Mara (UiTM) and Ministry of Higher Education, Malaysia for sponsoring the first author to pursue his study at the University of Nottingham (UoN), United Kingdom.

## REFERENCES

- [1] U. Bodenhofer and P. Bauer, “A Formal Model of Interpretability of Linguistic Variables,” in *Interpretability issues in fuzzy modeling*. Springer Berlin Heidelberg, 2003, pp. 524–545.
- [2] D. Nauck and R. Kruse, “Obtaining interpretable fuzzy classification rules from medical data,” *Artif. Intell. Med.*, vol. 16, no. 2, pp. 149–169, 1999.
- [3] R. Mikut, J. Jäkel, and L. Gröll, “Interpretability issues in data-based learning of fuzzy systems,” *Fuzzy Sets Syst.*, vol. 150, pp. 179–197, 2005.
- [4] L. A. Zadeh, “Outline of a New Approach to the Analysis of Complex Systems and Decision Processes,” *IEEE Trans. Syst., Man, and Cybern.*, vol. SMC-3, no. 1, pp. 28–44, 1973.
- [5] S.-M. Zhou, J. Garibaldi, R. John, and F. Chiclana, “On Constructing Parsimonious Type-2 Fuzzy Logic Systems via Influential Rule Selection,” *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 3, pp. 654–667, 6 2009.
- [6] Y. Jin, “Fuzzy modeling of high-dimensional systems: complexity reduction and interpretability improvement,” *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 2, pp. 212–221, 4 2000.
- [7] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, “Selecting fuzzy if-then rules for classification problems using genetic algorithms,” *IEEE Trans. Fuzzy Syst.*, vol. 3, no. 3, pp. 260–270, 1995.
- [8] O. Cordon, M. J. Del Jesus, F. Herrera, L. Magdalena, and P. Villar, “A Multiobjective Genetic Learning Process for joint Feature Selection and Granularity and Contexts Learning in Fuzzy Rule-Based Classification Systems,” in *Interpretability issues in fuzzy modeling*. Springer Berlin Heidelberg, 2003, pp. 79–99.
- [9] L. Koczy and K. Hirota, “Size reduction by interpolation in fuzzy rule bases,” *IEEE Trans. Syst. Man Cybern. - Part B*, vol. 27, no. 1, pp. 14–25, 1997.
- [10] Q. Liang and J. M. Mendel, “Designing interval type-2 fuzzy logic systems using an SVD-QR method: Rule reduction,” *Int. J. Intell. Syst.*, vol. 15, no. 10, pp. 939–957, 10 2000.
- [11] Y. Jin, “Decentralized adaptive fuzzy control of robot manipulators,” *IEEE Trans. Syst. Man Cybern. - Part B*, vol. 28, no. 1, pp. 47–57, 1998.
- [12] S.-M. Zhou and J. Gan, “Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modelling,” *Fuzzy Sets Syst.*, vol. 159, no. 23, pp. 3091–3131, 12 2008.
- [13] G. V. S. Raju, J. Zhou, and R. Kisner, “Hierarchical fuzzy control,” *Int. J. Contr.*, vol. 54, no. 5, pp. 1201–1216, 11 1991.
- [14] X.-J. Zeng and J. Keane, “Approximation Capabilities of Hierarchical Fuzzy Systems,” *IEEE Trans. Fuzzy Syst.*, vol. 13, no. 5, pp. 659–672, 10 2005.
- [15] R. Alcalá, J. Ramón Cano, O. Cordon, F. Herrera, P. Villar, and I. Zwir, “Linguistic modeling with hierarchical systems of weighted linguistic rules,” *Int. J. Approx. Reas.*, vol. 32, no. 2-3, pp. 187–215, 2 2003.
- [16] H. Ishibuchi, Y. Kaisho, and Y. Nojima, “Complexity, interpretability and explanation capability of fuzzy rule-based classifiers,” in *Proc. Int. Conf. Fuzzy Systems*. IEEE, 8 2009, pp. 1730–1735.
- [17] J. Casillas, O. Cordon, F. Herrera, and L. Magdalena, “Interpretability Improvements to Find the Balance Interpretability-Accuracy in Fuzzy Modeling: An Overview,” in *Interpretability issues in fuzzy modeling*. Springer Berlin Heidelberg, 2003, pp. 3–22.
- [18] K. Balazs and L. T. Koczy, “New parameterizable search space narrowing technique for adjusting between accuracy and interpretability in fuzzy systems,” in *Proc. IEEE 13th Int. Symp. Comput. Intell. and Inform. (CINTI)*. IEEE, 11 2012, pp. 323–328.
- [19] D. Nauck, “Measuring interpretability in rule-based classification systems,” in *Proc. 12th IEEE Int. Conf. Fuzzy Syst.* IEEE, 2003, pp. 196–201.
- [20] J. M. Alonso, S. Guillaume, and L. Magdalena, “A hierarchical fuzzy system for assessing interpretability of linguistic knowledge bases in classification problems,” in *Proc. PIMU*, 2006, pp. 348–355.
- [21] C. Mencar, C. Castiello, R. Cannone, and A. M. Fanelli, “Interpretability assessment of fuzzy knowledge bases: A cointension based approach,” *Int. J. Approx. Reas.*, vol. 52, pp. 501–518, 2011.
- [22] M. J. Gacto, R. Alcalá, and F. Herrera, “Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures,” *Inf. Sci.*, vol. 181, no. 20, pp. 4340–4360, 2011.
- [23] D. Pancho, J. Alonso, O. Cordon, A. Quirin, and L. Magdalena, “FINGRAMS: Visual Representations of Fuzzy Rule-Based Inference for Expert Analysis of Comprehensibility,” *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 6, pp. 1133–1149, 12 2013.
- [24] T. R. Razak, J. M. Garibaldi, C. Wagner, A. Pourabdollah, and D. Soria, “Interpretability indices for hierarchical fuzzy systems,” in *Proc. IEEE Int. Conf. Fuzzy Systems (FUZZ-IEEE)*. IEEE, 7 2017, pp. 1–6.
- [25] H. Ishibuchi and Y. Nojima, “Analysis of interpretability-accuracy trade-off of fuzzy systems by multiobjective fuzzy genetics-based machine learning,” *Int. J. Approx. Reas.*, vol. 44, no. 1, pp. 4–31, 1 2007.
- [26] M. Delgado, F. Von Zuben, and F. Gomide, “Multi-objective decision making: towards improvement of accuracy, interpretability and design autonomy in hierarchical genetic fuzzy systems,” in *Proc. Int. Conf. FUZZ-IEEE*, vol. 2. IEEE, 2002, pp. 1222–1227.
- [27] X. Zhu, J. Li, D. Wu, H. Wang, and C. Liang, “Balancing accuracy, complexity and interpretability in consumer credit decision making: A C-TOPSIS classification approach,” *Knowledge-Based Systems*, vol. 52, pp. 258–267, 11 2013.
- [28] H. Liu, A. Gegov, and M. Cocea, “Complexity Control in Rule Based Models for Classification in Machine Learning Context,” in *Advan. in Computat. Intell. Syst.* Springer, Cham, 2017, pp. 125–143.
- [29] M. Setnes and R. Babuska, “Rule base reduction: some comments on the use of orthogonal transforms,” *IEEE Trans. Syst., Man, and Cybern. - Part C*, vol. 31, no. 2, pp. 199–206, 5 2001.
- [30] J. Yen and L. Wang, “Simplifying fuzzy rule-based models using orthogonal transformation methods,” *IEEE Trans. Syst., Man, and Cybern. - Part B*, vol. 29, no. 1, pp. 13–24, 1999.
- [31] A. D. Benítez and J. Casillas, “Multi-objective genetic learning of serial hierarchical fuzzy systems for large-scale problems,” *Soft Comput.*, vol. 17, no. 1, pp. 165–194, 2013.
- [32] C. Mencar and A. Fanelli, “Interpretability constraints for fuzzy information granulation,” *Inf. Sci.*, vol. 178, no. 24, pp. 4585–4618, 2008.
- [33] J. M. Alonso, L. Magdalena, and G. González-Rodríguez, “Looking for a good fuzzy system interpretability index: An experimental approach,” *Int. J. Approx. Reas.*, vol. 51, no. 1, pp. 115–134, 12 2009.
- [34] S. Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956.